

Helmke, Andreas

## **Kommentar: Unterrichtsqualität und Unterrichtsklima: Perspektiven und Sackgassen**

*Unterrichtswissenschaft 30 (2002) 3, S. 261-277*



Quellenangabe/ Reference:

Helmke, Andreas: Kommentar: Unterrichtsqualität und Unterrichtsklima: Perspektiven und Sackgassen  
- In: Unterrichtswissenschaft 30 (2002) 3, S. 261-277 - URN: urn:nbn:de:0111-opus-76894 - DOI:  
10.25656/01:7689

<https://nbn-resolving.org/urn:nbn:de:0111-opus-76894>

<https://doi.org/10.25656/01:7689>

in Kooperation mit / in cooperation with:

# **BELTZ JUVENTA**

<http://www.juventa.de>

### **Nutzungsbedingungen**

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.  
Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

### **Terms of use**

We grant a non-exclusive, non-transferable, individual and limited right to using this document.  
This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

### **Kontakt / Contact:**

peDOCS  
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation  
Informationszentrum (IZ) Bildung  
E-Mail: [pedocs@dipf.de](mailto:pedocs@dipf.de)  
Internet: [www.pedocs.de](http://www.pedocs.de)

Digitalisiert

Mitglied der

  
Leibniz-Gemeinschaft

---

# Unterrichtswissenschaft

Zeitschrift für Lernforschung  
30. Jahrgang / 2002 / Heft 3

---

## Thema: Unterrichtsqualität

70 + 1.100

Verantwortlicher Herausgeber:  
Wolfgang Einsiedler

Wolfgang Einsiedler: Editorial: Das Konzept Unterrichtsqualität	194
Hartmut Ditton: Unterrichtsqualität – Konzeptionen, methodische Überlegungen und Perspektiven	197
Ferdinand Eder: Unterrichtsklima und Unterrichtsqualität	213
Hans-Günther Roßbach: Unterrichtsqualität im 2. Schuljahr – Ergebnisse einer empirischen Untersuchung	230
Marten Clausen, Kai Schnabel, Sabine Schröder: Konstrukte der Unterrichtsqualität im Expertenurteil	246
Andreas Helmke: Kommentar: Unterrichtsqualität und Unterrichtsklima – Perspektiven und Sackgassen	261

## Allgemeiner Teil

Jens Möller: Informationsbedingte Veränderungen elterlicher Einstellungen zu Berichtszeugnissen: eine experimentelle Studie	278
---	-----

---

Andreas Helmke

## **Kommentar: Unterrichtsqualität und Unterrichtsklima: Perspektiven und Sackgassen**

---

Der TIMSS-Schock über das erwartungswidrig mittelmäßige Abschneiden der deutschen Schülerinnen und Schüler bei der Third International Mathematics and Science Study hat zu einem bemerkenswerten, fast revolutionär zu nennenden Umschwung in der Bildungspolitik geführt, der heute als „empirische Wende“ bezeichnet wird (Lange, 2001). Die Umorientierung von einer input- und ressourcenorientierten Sichtweise (Verbesserung von Schule und Unterricht durch Steuerung der finanziellen, personellen, curricularen Rahmenbedingungen) zu einer „output“-orientierten Perspektive ist - nachdem PISA 2000 einen weiteren, mindestens ebenso großen Schock erzeugt hat - nicht mehr rückgängig zu machen. Was noch vor zehn Jahren für viele unerhört gewesen wäre, beruht heute auf einem großen Konsens: Schule muss sich gegenüber der Gesellschaft durch ihre nachweislichen Wirkungen und Effekte legitimieren. Diese Sichtweise hat zu einem unerhörten Aufschwung evaluativer Vorhaben aller Art geführt oder solche Unternehmungen zumindest erleichtert und ihnen Rückenwind gegeben: Von den großen internationalen Studien wie PISA über DESI<sup>1</sup>, über die drei großen Lernstandserhebungen in Hamburg (LAU), Brandenburg (QuaSÜM) und Rheinland-Pfalz (MARKUS) bis hin zu landeseinheitlichen Klassenarbeiten in mehreren Bundesländern.

Zugleich wächst die Erkenntnis, dass Evaluation allein nicht ausreicht. Jede Evaluation hat zwei Funktionen: einerseits Standortbestimmung und Rechenschaftslegung („system-monitoring“), andererseits Verbesserung („improvement“). Damit wird der Blick in zunehmendem Maße darauf gelenkt, was Schule leisten kann und muss, um beispielsweise die bei TIMSS und PISA zutage getretenen Defizite und Schwächen in Zukunft zu verringern. Innerhalb der Schulentwicklung ist es die *Entwicklung des Unterrichts*, die dabei im Mittelpunkt steht. Durch diese Rückbesinnung auf den Unterricht ist zugleich eine neue Forschungssituation entstanden, die eine zusammenfassende Sichtweise wichtiger Aspekte der Forschung zur Unterrichtsquali-

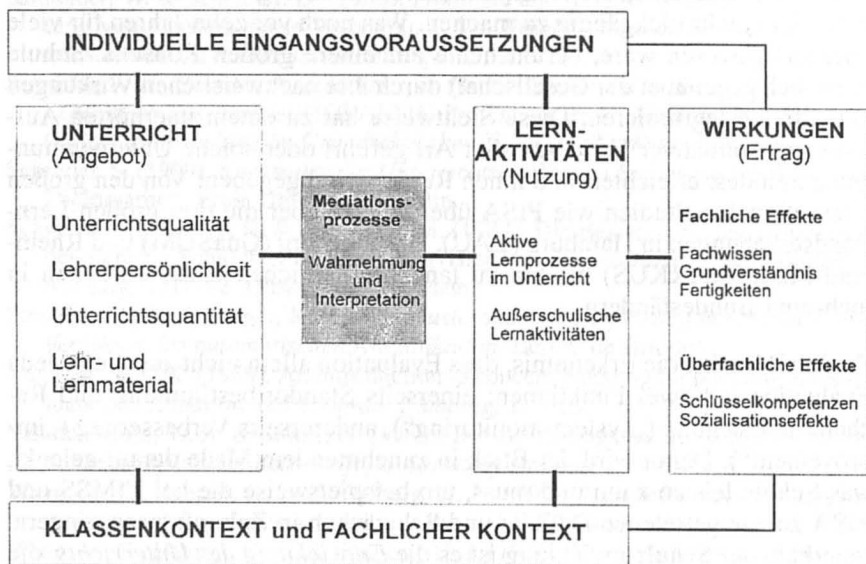
---

<sup>1</sup> Deutsch Englisch Schülerleistungen International: Projekt der Kultusministerkonferenz, das vom DIPF Frankfurt organisiert wird. Der Autor ist in dem interdisziplinär besetzten Konsortium federführend für die Bereiche Unterricht/Lehrer und Individuelle Bedingungen schulischer Leistungen.

tät, wie sie im vorliegenden Schwerpunktheft vorgenommen wird, besonders wichtig und lohnenswert erscheinen lässt. Ich werde im Folgenden versuchen, die in den vier Aufsätzen von Ditton, Eder, Roßbach sowie Clausen, Schnabel & Schröder berichteten empirischen Ergebnisse und Standpunkte aufzugreifen und an einigen Stellen mit eigenen Sichtweisen zu ergänzen.

Um einen gemeinsamen Bezugspunkt für alle vier Beiträge herzustellen, möchte ich mit der folgenden Abbildung wichtige Aspekte des Unterrichts und seiner Wirkungsweise in Form eines Modells darstellen. Dieses Modell soll zugleich demonstrieren, dass meine Einschätzung der hier vorliegenden Beiträge keineswegs voraussetzungslos ist, sondern an bestimmten Positionen festgemacht wird, die zu einem Teil im folgenden Rahmenmodell veranschaulicht werden. Es handelt sich um die Kurzfassung einer differenzierten Darstellung bei Helmke (in Druck). Dem Modell liegen Überlegungen von Fend (1998) sowie von Helmke & Weinert (1997) zugrunde.

Abbildung 1:  
Ein erweitertes Angebots-Nutzungs-Modell des Unterrichts



Mit diesem Modell sollen folgende Punkte veranschaulicht werden:

- Unterricht ist in seiner Gesamtheit ein *Angebot*, das je nach individuellen Lernaktivitäten mehr oder weniger *genutzt* werden kann. Auf der Angebotsseite lassen sich vier große Erklärungsblöcke unterscheiden (in Klammern jeweils beispielhaft wichtige Teilaspekte): *Unterrichtsqualität* (Klassenführung, Klarheit, Passung, Methodenvielfalt, Motivierung, Wertschätzung der Schüler/innen, angemessene Leistungsrückmel-

dung); *Lehrerpersönlichkeit und -expertise* (implizite Theorien, Erwartungen, pädagogische Orientierungen und Ziele, fachliche Expertise, diagnostische Kompetenz, Selbstwirksamkeit, Engagement, Geduld sowie eigene Lern- und Reflexionsbereitschaft); *Unterrichtsquantität* (nominale Stundentafel, Unterrichtsausfall, Schülerfehlzeiten, Zeitnutzung, Lerngelegenheiten/Content Covered) und die *Qualität des Lehr- und Lernmaterials*.

- Von großer Bedeutung sowohl für den Unterricht als auch für die unterrichtlichen Wirkprozesse sind der *Klassenkontext* und der *fachliche Kontext*. Der erstere umfasst sowohl Aspekte der Klassenzusammensetzung (z.B. Anteil von Schüler/innen aus bildungsfernen Schichten und mit Deutsch als Mutter- oder Fremdsprache) als auch den normativen Kontext sowie Aspekte des Schüler-Schüler-Verhältnisses und des Lehrer-Schüler-Verhältnisses. Gelegentlich spricht man bei diesen Konstrukten auch von „Klima“. Je nach fachlichem (und damit zugleich curricularem Kontext) ist mit qualitativen Unterschieden der Wirkungszusammenhänge zu rechnen: Was für Mathematik/5. Klassenstufe gilt, lässt sich nur begrenzt auf andere Fächer und/oder Klassenstufen übertragen.
- „Den“ guten Unterricht kann es schon deshalb nicht geben, weil je nach angepeiltem *Lernziel* (Aufbau von Fertigkeiten; Förderung von Grundverständnis; Erzeugung intelligenten Wissens; Entwicklung sozialer Kompetenzen oder methodischer Fähigkeiten) ganz unterschiedliche Lehrmethoden angebracht sind.
- Dazu kommt, dass Unterschiede in den *Eingangsvoraussetzungen* der Schülerinnen und Schüler sowie Unterschiede in der *Klassenzusammensetzung* ebenfalls zu einer Relativierung des „guten“ Unterrichts führen: Ein straffer, direkter, stark lehrergesteuerter, strukturierter Unterricht wird bei ängstlichen und leistungsschwachen Schülerinnen und Schülern und im Fach Mathematik angemessen sein, in anderen Fällen und Fächern kann er „kontraindiziert“ sein. Und die optimale Dosierung von Klassenführung, Sozialpädagogik und Instruktion dürfte in einer Schule im sozialen Brennpunkt, wo Deutsch überwiegend Fremdsprache ist, anders aussehen als etwa in homogenen Klassen oder Kursen im ländlichen Raum.

Nun zu den vier Beiträgen im Einzelnen:

## DITTON

Ditton gibt einen kurzen und prägnanten Überblick über Konzeptionen der Unterrichtsqualität und verknüpft dies mit einer Übersicht über damit zusammenhängende wichtige Methodenfragen. Dies ist ein komplexes Unterfangen, mit dem andere ganze Bücher gefüllt haben, und in Anbetracht des zur Verfügung stehenden Platzes ist der Versuch aus meiner Sicht als geglückt zu

bezeichnen. Es gelingt Ditton, einige Schlüsselfragen anzusprechen und Antworten zu skizzieren, die im Bereich der Unterrichtsqualität notwendigerweise auftauchen.

### *Prozess- versus produktorientierte Perspektive*

Gleich zu Beginn macht Ditton auf zwei grundlegend verschiedene Sichtweisen der Unterrichtsqualität aufmerksam, die in der allgemeinen Diskussion nicht immer klar getrennt werden: Entweder der *Prozess* des Unterrichtens - seine Methodik, die Lehrer-Schüler-Interaktion, die Klassenführung etc. - ist selbst Gegenstand der Beurteilung, oder die Qualität des Unterrichts bemisst sich an seinen *Wirkungen*. Im ersteren Falle sind normative Vorstellungen über ge- oder misslungenen Unterricht nötig (vgl. hierzu den folgenden Abschnitt). Damit ist ein wichtiger Punkt angesprochen, über den zwischen geisteswissenschaftlich orientierten Vertretern der Lehrerbildung einerseits und empirisch orientierten Erziehungswissenschaftlern und Bildungsforschern andererseits häufig Dissens besteht. Allerdings hat die ausführliche (und teilweise sehr kontroverse) Diskussion nach TIMSS und erst recht die nach PISA 2000 (sowie nach den großen Länderstudien LAU, QuaSUM und MARKUS) zu der anfangs beschriebenen Wende im Denken geführt: dass sich nämlich Schule und Unterricht mit Sicherheit nach ihren Wirkungen messen lassen müssen, und dies hat Folgen für die Frage nach der Unterrichtsqualität.

Das heißt selbstverständlich nicht, dass die Frage nach dem „guten Unterricht“ - also all das, was einen guten Teil der Lehrerbildung ausmacht oder ausmachen sollte - obsolet geworden wäre. Dies ist keineswegs der Fall. Jedoch haben TIMSS, PISA etc. im Bereich der traditionellen Lehrerbildung, speziell der Allgemeinen Didaktik, für eine gehörige Portion frischer empirischer Luft gesorgt. Ich stimme hier Baumert & Köller (2000) vollkommen zu, wenn sie resümieren: „Für die allgemeine Didaktik heißt dies, Abschied zu nehmen von normativen, wertrational begründeten Unterrichtsbildern, die gegen empirische Überprüfung immunisiert sind, und für die Unterrichtsforschung, die Vorstellung einer instrumentell-technologischen Gestaltbarkeit von Unterricht durch die Manipulation isolierter Maßnahmen endgültig aufzugeben“ (S. 271).

Zu den Sackgassen einer auf Methoden fixierten Konzeption der Unterrichtsqualität möchte ich Weinert (1998) zitieren:

*„Welch ein Glaube an die Zerlegbarkeit der Lehrstoffe in Lektionen, welche Vision von der stereotypen Einheitlichkeit der psychologischen Lernmechanismen und welche Hoffnung auf die unbegrenzten Möglichkeiten des Lehrens als Lernen! Kein Wunder, dass der Kampf gegen die Formalstufen des Unterrichts zur großen Gemeinsamkeit der sonst sehr zerstrittenen Reformpädagogen wurde.“*

*Methoden zu Bewertungskriterien guten Unterrichts zu machen, ist aber kein Relikt des 19. Jahrhunderts, sondern findet sich auch in der Gegenwart in vielfältiger, oft versteckter, sich modern gebender, meist aber ideologisch eingefärbter Weise. Unterricht - so heißt es in einem kakophonem Chor - sei nur dann gut, wenn er lebenspraktische Projekte zum Inhalt hat, wenn er als Teamarbeit organisiert ist, wenn er völlig offen für die Gestaltung durch Schüler bleibt oder wenn er statt wichtiger Lerninhalte ausschließlich das Lernen lehrt. Solchen radikal einseitigen Methodenfixierungen - die z.B. im Erstleseunterricht lange Zeit zu ebenso erbitterten wie unsinnigen didaktischen Kontroversen geführt haben - wird im Folgenden deutlich widersprochen. Methoden sind nicht Selbstzweck des Unterrichts, sondern Werkzeuge zur Erreichung bestimmter Ziele und nur als solche brauchbar oder unbrauchbar, gut oder schlecht“ (S. 8).*

### *Allgemeine Didaktik und Forschung zur Unterrichtsqualität*

Besonders wichtig und verdienstvoll erscheint mir in Dittons Artikel zu sein, dass auch ganz explizit Modelle und Gedanken der Allgemeinen Didaktik herangezogen und mit gebührender Ausführlichkeit berichtet werden. Dies zu unterlassen ist eine häufige Schwäche vieler pädagogisch-psychologisch orientierter empirischer Arbeiten, die nicht über den Tellerrand ihrer eigenen Disziplin hinausschauen. Und wer sich die diesbezüglich defizitären Artikel in den einschlägigen angloamerikanischen Handbüchern und Enzyklopädien zum Thema „Teaching“ anschaut, wird den in der Didaktik unternommenen Versuch einer bildungstheoretischen Fundierung als besonders wichtig, ja nötig anerkennen.

Terhart (2001) hat das Verhältnis zwischen Lehr-Lern-Forschung und Pädagogischer Psychologie einerseits und Didaktik andererseits meines Erachtens sehr treffend als das Verhältnis zweier fremder Schwestern bezeichnet und eine Reihe von Gründen für diese Fremdheit benannt. Einer davon ist die überwiegend empirische Orientierung der Pädagogischen Psychologie im Gegensatz zur eher normativ orientierten Didaktik - und in der Tat sind Abstraktionsniveau und Duktus didaktischer Texte oft auf einem Niveau, das den Bezug zu empirischen Studien nicht unbedingt nahe legt. Ein zweites Handikap, das Ditton anspricht, ist die begrenzte Nützlichkeit solcher sehr allgemeinen, in keiner Weise fachwissenschaftlichen oder fachdidaktischen Modelle für die Lehrerbildung. Hier könnten - zum beiderseitigen Nutzen - noch viele wichtige Brücken gebaut werden: empirische Anreicherung der Allgemeinen Didaktik und bildungstheoretische Absicherung der empir(ist)ischen Unterrichtsforschung.

### *Allgemeine Unterrichtsforschung vs. Fachdidaktik*

Auch an diesen von Ditton zu Recht als wichtig dargestellten Punkt möchte ich anknüpfen: Die Diskussion nach der Publikation der TIMS-Studie und

die jetzt nach PISA 2000 einsetzenden Überlegungen zeigen immer deutlicher, dass eine bereichsunspezifische Sichtweise des Unterrichts und seiner Qualität, die die spezifische Eigenart des Faches, den Aufbau seines Curriculums, die sehr unterschiedlichen Fachdidaktiken und die divergierenden Lernziele ganz außer Acht lässt, nicht mehr zeitgemäß ist. Dieser kritische Einwand trifft übrigens auch meine eigenen vergangenen Studien (z.B. die mit Weinert gemeinsam organisierte Münchner Studie, die den deutschen Teil der internationalen Classroom Environment Study der IEA repräsentiert). Obwohl auch damals innerhalb der IEA bereits fundierte fachdidaktische Diskussionen stattfanden, die in der FIMS und der SIMS (First bzw. Second International Mathematics Study) teilweise ihren Niederschlag fanden, hat man die Frage des Inhaltsbezuges des Unterrichts vor 20 Jahren noch nicht als so entscheidend beurteilt. Zu den wenigen, die dies - lange vor TIMSS - bereits damals kritisch angemahnt haben, gehört Achtenhagen. TIMSS hat sich als Katalysator herausgestellt, um fachdidaktische und erziehungswissenschaftliche Expertise in Form gemeinsamer Projekte, Modelle und Instrumente im Bereich der Mathematik und Naturwissenschaften zu bündeln und synergetisch zusammenzuführen. Zu einem wesentlichen Teil ist dies das Verdienst von J. Baumert, der seinerzeit - als Direktor des IPN Kiel - solche Brücken gebaut und forciert hat. Im Hinblick auf die noch nicht abgeschlossenen bzw. gerade erst begonnenen großen Evaluationsprojekte DESI und PIRLS / IGLU bleibt zu hoffen, dass dies für den Bereich des Deutsch- und Fremdsprachenunterrichts bzw. für den Unterricht in der Grundschule in gleicher Weise gelingen möge.

### *Schulqualität und Unterrichtsqualität*

Ein wichtiger inhaltlicher Punkt ist das Verhältnis von Schulqualitätsforschung und Forschung zur Unterrichtsqualität. Ditton selbst hat zu dieser Frage viel geforscht und gilt zugleich als Experte für mehrebenenanalytische Verfahren. Er hat immer wieder zu Recht darauf hingewiesen, dass die Suche nach der „guten Schule“ - was ihre fachlichen und überfachlichen Effekte anbelangt - in mancher Hinsicht eine Sackgasse darstellt. Ohne eine simultane Untersuchung sowohl der Schul- als auch der Schulklassenebene gelangt man sehr leicht zu Artefakten. Alle mir bekannten neueren mehrebenenanalytisch angelegten Studien (z.B. Arnold, in Druck; Hosenfeld, Helmke, Ridder & Schrader, 2001) haben gezeigt, dass die durch Schulklassenunterschiede erklärte Varianz der Schülerleistungen deutlich größer ist als diejenige der Schulunterschiede. Bei aller unbestrittenen Wichtigkeit der mit zunehmender Schulautonomie verbundenen Maßnahmen und Programme (wie Schulprofil, schulinterne Lehrerfortbildung, Schulgeist, Schulbroschüre etc.) scheint mir bei manchen Vertretern der Schulentwicklung der Blick auf den eigentlichen Grundgedanken der Schule etwas abhanden gekommen zu sein: Schulentwicklung muss im Kern Unterrichtsentwicklung sein. Die Berechnung von „Schulmittelwerten“ zu Maßen der Unterrichtsqualität oder der Effizienz der Klassenführung kann völlig irreführend sein - wenn sich



nämlich Klassenzusammensetzung, Unterrichtsqualität und Lehrer-Schüler-Interaktion zwischen den Klassen ein und derselben Schule voneinander unterscheiden. Und genau das ist sehr häufig der Fall.

## CLAUSEN, SCHNABEL & SCHRÖDER

Welches sind die Schlüsseldimensionen der Unterrichtsqualität, welche die zentralen Komponenten und wichtigsten Lehrerkompetenzen? Auf diese klassischen Fragen der Unterrichtsforschung gibt es inzwischen eine kaum noch übersehbare Zahl von Antworten. Angesichts der Flut von Konzepten aus der Unterrichtsforschung, der Lehr-Lern-Forschung und der Schul- und Unterrichtsklimaforschung sind solche Versuche der Systematisierung unabdingbar, um eine Orientierung zu ermöglichen. Um den Stellenwert der Arbeit von Clausen et al. besser zu erläutern, soll zunächst skizziert werden, dass es mehrere sehr unterschiedliche Ansätze gibt, das Chaos der unterrichtsrelevanten Konstrukte in eine Ordnung zu bringen.

### *Theoretische Ansätze*

Solche Orientierungen können zunächst theoretisch erfolgen, und zwar in zweierlei Weise: a) *lehr- bzw. kompetenzorientiert* oder b) *lernorientiert*. Zum erstgenannten Ansatz gehören die Rahmenmodelle von Komponenten der Unterrichtsqualität und Unterrichtswirkungen (vgl. Abbildung 1), die auf einer Sichtung und Bewertung des Forschungsstandes basieren, wie er von Unterrichtsforschern wie Weinert, Brophy oder Slavin vorgenommen wird. Weinert unterscheidet beispielsweise vier zentrale Kompetenzen von Lehrern: didaktische, klassenführungsbezogene, diagnostische und fachinhaltliche Kompetenzen. Slavin hat die Vielfalt von Konstrukten der Unterrichtsqualität in Gestalt seines QuAIT-Modells ebenfalls auf vier Aspekte reduziert: *Quality of instruction*, *Appropriateness*, *Incentives* und *Time*. Dies sind alles Ansätze, die von einer Makroebene des Unterrichts ausgehen.

Demgegenüber stehen Ansätze, die von den Prozessen des Lernens ausgehen und daraus ableiten, welche Qualitätsmerkmale Prozesse der Instruktion besitzen müssen. Zu diesen Modellen kann man beispielsweise das Modell der Stationen des Lernprozesses von Gagné und Driscoll (1988) zählen, dessen acht Stufen des Lernprozesses korrespondierende unterrichtliche Ereignisse gegenübergestellt werden. Abbildung 2 zeigt dieses Modell.

Vom Denkprinzip her ähnlich ist das Modell von Shuell (1996). Ich halte diese an Stationen des Lernprozesses orientierten Ansätze für ein Verständnis der Wirksamkeit verschiedener Aspekte des Lehrens und Lernens für außerordentlich wichtig und nützlich. Soweit zwei theoretisch orientierte Ansätze.

**Abbildung 2:**  
**Stationen des Lernprozesses nach Gagné & Driscoll, 1988**

Learning Process	Instructional Event	
<b>Attention: Alertness</b> <i>Aufmerksamkeit, Wachheit</i>	<b>Gaining Attention</b> <i>Aufmerksamkeit herstellen</i>	1
<b>Expectancy</b> <i>Erwartungen</i>	<b>Informing Learner of the Objective</b> <i>Orientierung, Motivierung</i>	2
<b>Retrieval to Working Memory</b> <i>Abruf von Wissen aus dem Langzeit- ins Arbeitsgedächtnis</i>	<b>Stimulating recall of prior knowledge</b> <i>Vorwissensnutzung anregen</i>	3
<b>Selective Perception</b> <i>Selektive Verarbeitung</i>	<b>Presenting the Stimulus Material</b> <i>Präsentation d. Lernmaterials</i>	4
<b>Encoding: Entry to LTM Storage</b> <i>Speichern im Langzeitgedächtnis</i>	<b>Providing Learning Guidance</b> <i>Unterstützung des Lernens</i>	5
<b>Responding</b> <i>Offenes Lernverhalten</i>	<b>Eliciting Performance</b> <i>Üben und Probieren</i>	6
<b>Reinforcement</b> <i>Verstärkung</i>	<b>Providing Feedback</b> <b>Assessing Performance</b> <i>Feedback und Korrekturen</i>	7 8
<b>Cueing Retrieval</b> <i>Abruf und Nutzung von Wissen aus dem Langzeitgedächtnis</i>	<b>Enhancing Retention + Transfer</b> <i>Verbesserung von Behalten und Transfer</i>	9

### *Empirische Ansätze*

Geht man empirisch vor, so lassen sich ebenfalls völlig unterschiedliche Strategien verfolgen. Ein bekannter Ansatz ist derjenige der *Meta-Analyse*, deren Prinzip die Zusammenfassung (in Form von Kennwerten) vieler unterschiedlicher Studien zum gleichen Gegenstand ist. Die wichtigsten Meta-Analysen zur Rolle des Unterrichts stammen aus der Gruppe um Walberg (z.B. Wang et al., 1993); und nicht zufällig ist es neben Brophy auch Walberg, der gemeinsam mit Paik eine einflussreiche und im Internet verfügbare Publikation der International Academy of Education verfasst hat (Walberg & Paik, 2000).

Geht es nicht um die Effektbestimmung, sondern um die Dimensionalität unterrichtlicher Faktoren, so liegt die Faktorenanalyse nahe, die an dieser Stelle nicht weiter kommentiert werden muss. Die *Faktorenanalyse* gehört zu den variablenzentrierten Verfahren, d.h. die Reduktion der Variablenvielfalt erfolgt aufgrund der korrelativen Muster zwischen Variablen - ohne Berücksichtigung ihres Auftretens und ihres Mischungsverhältnisses bei konkreten Personen.

Clausen et al. verfolgen mit ihrer Untersuchung ebenfalls einen variablenzentrierten Ansatz: Ausgehend von den empirisch vorgefundenen Klassifikationen und Unterscheidungen (durch Sortieren) analysieren sie das Ergebnismuster, um so Aussagen über zentrale Koordinaten zu erhalten.

Demgegenüber stehen *personenzentrierte* Analysestrategien: In einem ersten Schritt werden z.B. besonders erfolgreiche (oder erfolglose) Lehrpersonen (Experten, Meisterlehrer etc.) identifiziert, wobei das Kriterium sowohl prozessorientiert sein kann (Unterrichtsbeurteilung durch die Schulaufsicht oder im Rahmen eines Forschungsprojektes) als auch produktorientiert (Lehrkräfte mit nachweislich positiven Entwicklungen ihrer Klasse, z.B. überdurchschnittlicher Leistungszuwachs oder Angstabbau). Im zweiten Schritt werden Unterricht und Klassenkontext dieser Extremgruppen beschrieben und analysiert, um Anhaltspunkte für den Erfolg dieser Personen zu suchen.

### *Ergebnisse der MDS-Analyse*

Clausen et al. gehen mit Hilfe der „multidimensionalen Skalierung“ der Frage nach, welche nach der Einschätzung von Experten die übergreifenden, die wichtigsten, die real unterscheidbaren Komponenten und Dimensionen der Unterrichtsqualität sind. Dies ist ein origineller und relativ selten gewählter Ansatz. Das Ergebnis ist bemerkenswert, und zwar in mehrfacher Hinsicht: Die „Wolke“ der MDS-skalierten Begriffe lässt sich offenbar ganz gut durch zwei Dimensionen beschreiben. Das horizontale Kontinuum reicht vom *Klassenklima* (Aspekte des Verhältnisses der Schülerinnen und Schüler untereinander) bis zu *didaktischen Aspekten der Instruktion*, und das vertikale Kontinuum wird durch die Endpunkte „effektiv leistungsorientierte Instruktion“ einerseits und „individualisierende sozialorientierte Instruktion“ andererseits aufgespannt. Interessant ist auch die Clusteranalyse der Merkmale, die zu den folgenden acht Bereichen führt: (1) Sozialorientierung, (2) Sozialklima, (3) Einstellungen und Lernhaltungen, (4) Didaktische Qualität, (5) Effizienz der Klassenführung, (6) Zeitnutzung, (7) Motivationsunterstützung und (8) Lehrerzentriertheit.

### *Ideen für Folgeuntersuchungen*

Es wäre interessant, die auf diese Weise erhaltene Liste einmal den Ordnungsansätzen und Faktoren der Unterrichtsqualität gegenüberzustellen, wie sie im gleichen Artikel eingangs erfolgte, insbesondere den Konzepten von Einsiedler und Wang et al. Zeigen sich in der empirischen Untersuchung von Clausen et al. in Form ihres Acht-Cluster-Ergebnisses Bereiche, die bei den klassischen Ansätzen (z.B. bei den von Ditton erwähnten Ansätzen wie Slavins QuAIT-Modell) nicht oder in anderer Gestalt (oder auf anderem Niveau) auftauchen? Das Umgekehrte dürfte eigentlich nicht der Fall sein,

wenn die 87 Ausgangsbegriffe, die die Autoren ihrer Studie zugrunde legen, wirklich „flächendeckend“ das gesamte Feld abgegrast haben.

Was mir bei der Durchsicht der 87 Konstrukte (die, auf Kärtchen aufgedruckt, sortiert werden sollten) aufgefallen ist (siehe Abbildung 3 bei Clausen et al.): dass ich selbst, hätte ich diese Aufgabe als Versuchsperson erledigen müssen, da und dort überfordert gewesen wäre. Was zum Beispiel hätte ich mit Konstrukten wie „Ausstattung“, „Bevorzugung“, „Gemeinschaft“, „Instrumentelle Aktivität“ angestellt: Ausstattung? (finanziell? materiell? Schule?); Bevorzugung? (von was oder wem?) etc. Oder sind die Labels in Tabelle 3 lediglich Kurzformen, stand auf den Kärtchen noch erläuternder Text? Das kann ich mir eigentlich nicht anders vorstellen. Das gleiche gilt für zahlreiche andere Konzepte, z.B. auch für „Bezugsnormorientierung“: welche Variante, die individuelle oder die soziale?

Möglicherweise wäre das Gesamtergebnis auch noch klarer - und den klassischen Schemata leichter zuzuordnen - gewesen, wenn man sich bei der Sammlung der zunächst 143, anschließend 87 verbliebenen Konstrukte (basierend auf den genannten Inventaren) wirklich auf *Unterricht* beschränkt und auf Aspekte des Klassenklimas (wie Cliquenbildung) und erst recht auf Facetten der Lernbereitschaft und der Mitarbeit (dies sind Bedingungen und zugleich Folgen des Unterrichts, aber es sind keine Unterrichtskonstrukte) verzichtet hätte.

Ich fände es ausgesprochen nützlich und gewinnbringend, auf dem von Clausen et al. beschrittenen Weg weiterzugehen und die Methode der Multidimensionalen Skalierung wesentlich stärker als bisher in der erziehungswissenschaftlichen Forschung üblich zu nutzen. Neben der bereits angesprochenen Konzentration auf einen Inhaltsbereich - nämlich Unterricht - sehe ich eine Reihe anderer Fragestellungen, die mit der gleichen Methode fruchtbar verfolgt werden könnten: *Anstelle der hier gewählten Konstrukte könnte man z.B. einen Schritt in Richtung Verhaltensnähe gehen und - etwa im Sinne der ebenfalls stark verhaltens- und kompetenzorientierten 88 Standards des Lehrerberufs von Oser (1999) - danach fragen, ob und wie diese Standards von Lehramtsstudierenden und Lehrkräften unterschieden werden.* Dies führt auf eine zweite Erweiterung, die ich ebenfalls interessant fände: *Anstelle der hier gewählten Experten (und der nur am Rande erwähnten Gruppe von Studierenden) könnte man aus einer entwicklungspsychologischen und kompetenztheoretisch orientierten Sicht fragen, ob der Verlauf des Studiums (z.B. des Lehramtes) bzw. die berufliche Sozialisation von Lehrkräften insbesondere in den ersten Berufsjahren mit systematischen Veränderungen der Qualität und Differenziertheit im Bereich unterrichtsrelevanter Konstrukte verbunden ist.* Dies müsste man mit einer Recherche der subjektiven Theorien (Alltagstheorien, epistemologische Theorien, intuitive Konzepte) und zum Lehrerwissen (Fischler, 2001) verbinden (siehe auch das Kapitel zum mathematischen Weltwissen im Band TIMSS/III von Baumert & Klieme).

## ROSSBACH

Diese Arbeit gehört zu den relativ wenigen, die der Frage nach der Unterrichtsqualität in der *Grundschule* empirisch nachgehen. Mir sind außer unseren eigenen Arbeiten im Projekt SCHOLASTIK (Helmke & Schrader, 1997; Weinert & Helmke, 1997) nur wenige empirische Arbeiten bekannt, die sich mit dem Unterricht in der Grundschule befassen. Als Teil der „European Child Care and Education Study“ war der deutsche Beitrag an internationale Vorgaben gebunden, wozu auch die Gestaltung der Stichprobe gehört: Die 272 Schülerinnen und Schüler verteilen sich auf 214 Schulklassen, sodass infolge dieser Restriktion zwar der Unterricht in diesen 214 Klassen beobachtet und beurteilt werden konnte, ohne jedoch diese Unterrichtsdaten mit klassenweise aggregierten Schülerdaten in Verbindung bringen zu können. Schülerperzeptionen des Unterrichts liegen also nicht vor und - wenn ich es richtig verstanden habe - Angaben der Schülerinnen und Schüler zu ihrem Befinden, ihrer Motivation offenbar ebenso wenig. Dies dem Autor vorzuwerfen, wäre unfair - es sind eben Restriktionen, an die die deutsche Teilstudie gebunden war, die allerdings den Interpretationsspielraum empfindlich einschränken.

Die Qualität des Unterrichts wird somit ausschließlich über die Unterrichtsbeurteilungen mit Hilfe von Rating-Skalen erfasst; zusätzlich pädagogische Orientierungen der Lehrkräfte in Form eines Interviews. Die faktorielle Struktur umfasst (a) Klassenmanagement, (b) Schülerorientierung, (c) Direktiver Unterricht, (d) Anwendungsbezug und (e) Unterstützung (bei Stillarbeit).

Wie Roßbach zu Recht sagt, ist eine einmalige zweistündige Unterrichtsbeobachtung in der Tat anfällig für Fehler verschiedenster Art - aber leider hat man als Unterrichtsforscher(in) oft keine andere Wahl. Schade, dass nicht zumindest in einem Teil der Klassen der Unterricht videografiert werden konnte - dies hätte die nachträgliche Beurteilung der Unterrichtsqualität durch verschiedene Beobachter und aus verschiedenen Perspektiven ermöglicht. Zur Erfassung der Beobachterübereinstimmung wurden in 21 Klassen zwei Beobachter zu verschiedenen Zeitpunkten eingesetzt. Die vorgefundene Übereinstimmung zwischen den Beobachtern erscheint ausreichend, wenn man keine unrealistisch hohen Maßstäbe anlegt. Allerdings wäre es ratsam gewesen, sich bei der Planung und Auswertung der Unterrichtsbeobachtung des Werkzeugs der *Generalisierungstheorie* zu bedienen, die für genau solche Zwecke geschaffen ist (Renkl & Helmke, 1993).

Weshalb die Unterrichtsbeurteilung auf der Basis des von uns (Helmke & Schrader, 1997) entwickelten Ratingbogens in dieser Studie so anders ausfällt, insbesondere als seinerzeit im Projekt SCHOLASTIK (das im Großraum München stattfand), ist auch mir ein Rätsel. Mangels entsprechender Anker wird sich dieses Rätsel wohl nie lösen lassen. Forschungsstrategisch lässt sich daraus für den Einsatz von Unterrichts-Ratings in künftigen Studien die Empfehlung ableiten, das Training der Beobachter(innen) bzw. Beurteiler(innen) in einer Weise vorzunehmen und zu dokumentieren (insbesondere durch die Verwendung geeigneter Video-Ausschnitte und zugehöri-

ger Beispiel-Kodierungen), dass ein Mindestmaß an Vergleichbarkeit über verschiedene Zeitpunkte, Regionen und Forschergruppen ermöglicht wird.

Zu den bemerkenswertesten Ergebnissen zählt für mich, dass die Klassenzusammensetzung (und auch die Orientierungen der Klassenlehrerinnen und Klassenlehrer) die Unterrichtsqualität in keiner Weise determiniert. Es ist plausibel und empirisch belegt, dass das Unterrichten je nach vorgefundenen Bedingungen (große vs. kleine Klasse, homogene vs. heterogene Schülerschaft, geringe vs. solide Vorkenntnisse, Deutsch überwiegend Erst- vs. Zweitsprache) erleichtert oder auch erschwert sein kann. Die Einschränkung der Unterrichtsqualität durch solche, von den Lehrkräften nicht beeinflussbaren Rahmenbedingungen wird jedoch häufig überschätzt. In dem soeben abgeschlossenen Projekt MARKUS<sup>2</sup>, bei dem der gesamte Schülerjahrgang der 8. Klassenstufe des Bundeslandes Rheinland-Pfalz in Mathematik getestet wurde, begleitet von Fragebögen an alle Schülerinnen und Schüler sowie an die beteiligten Lehrkräfte, kommen wir ebenfalls zu dem Schluss, dass der Klassenkontext zwar den Unterricht in seiner Gestaltungsmöglichkeit einschränkt, jedoch keine Fessel darstellt: Das wichtigste Einzelmerkmal (die durchschnittliche Bildungsferne der Eltern) erklärt maximal 10% der Leistungsstreuung zwischen Klassen, und die Summe aller Merkmale erklärt im Höchstfall 15% der Leistungsvarianz. Da bleibt noch sehr viel Spielraum für die unterrichtliche Gestaltung (Hosenfeld, Helmke, Ridder & Schrader, 2001).

Dass die auf den Ratings basierenden Unterrichtsskalen nur sehr schwach mit der Schulleistung zusammenhängen, ist auf den ersten Blick befremdlich. Schaut man sich in der einschlägigen Literatur um, dann sind auch dort oft nur sehr schwache Korrelationen zu finden. Ein Schwachpunkt ist hier natürlich, dass kein curricular valider Schulleistungstest entwickelt wurde, sondern dass aus naheliegenden Gründen eine international vergleichbare Testbatterie eingesetzt wurde, und zwar eine deutsche Adaptation des Woodcock Johnson Test of Achievement. Anstelle der in Roßbachs Tabelle 3 berichteten Korrelation der vier Unterrichtskennwerte mit „den Schulleistungen“ (also mit dem Gesamtscore dieses Leistungstests) könnte und sollte man diesen Gesamttestwert einmal dekomponieren, um festzustellen, ob sich für die Subscores (Satzverständnis, Rechnen, Problemlösen, Sachkunde) differentielle Ergebnisse zeigen. Weitere Erklärungen für die Unterschiede zwischen den hier gefundenen Ergebnissen und den bei SCHOLASTIK gefundenen (dort zeigten sich immerhin Korrelationen des Mathematik-Leistungszuwachses mit dem Unterricht zwischen 0,36 bei der Effizienz der Klassenführung und 0,34 für das soziale Klima) könnten darin liegen, dass es bei SCHOLASTIK um etwas ältere Schülerinnen und Schüler ging (3. und 4. Klasse). Auf der anderen Seite hatten wir dort Korrelationen mit Leistungszuwachsen

---

<sup>2</sup> *Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulqualität (Helmke & Jäger, in Druck; Helmke et al., 2002)*

berichtet, die üblicherweise noch geringer ausfallen als Korrelationen mit einfachen Mittelwerten.

## **EDER**

Welche Rolle spielen Schul-, Klassen- und Unterrichtsklima für die Unterrichtsqualität? Welchen theoretischen Stellenwert haben sie, wie können die Klimakonstrukte die Ziele der Unterrichtsforschung unterstützen, nämlich eine theoriegeleitete Beschreibung und Erklärung von Erscheinungsformen und Konsequenzen des Unterrichts mit dem Ziel sowohl der Bestandsaufnahme als auch der Optimierung von Lehr-Lern-Prozessen?

### *Mehrperspektivität der Erhebung des Unterrichts*

Es besteht kein Zweifel daran, dass es nötig und wichtig ist, die „Klientel“ - seien es die Schülerinnen und Schüler, seien es Studierende oder Lehrkräfte - nach ihrer Einschätzung unterrichtlicher Aspekte sowie nach ihrer Befindlichkeit zu fragen. Eines der Grundprinzipien der Unterrichtsforschung ist ja gerade das Erfordernis, sich bei der Erfassung des Unterrichts nicht auf eine Quelle, eine Gruppe von Adressaten und eine Methode zu verlassen, sondern ein breites Spektrum von Perspektiven und Methoden einzubeziehen (vgl. Helmke & Weinert, 1997). Jede dieser Quellen und Methoden hat ihre je eigenen Vorteile und Nachteile, ihre Stärken und ihre blinde Flecken. Clausen (2001) hat das in seiner Arbeit gut zeigen können. Die Mehrheit vorangegangener Arbeiten, die sich nicht nur mit Schülerfragebögen begnügt hat, sondern daneben auch die Sichtweise von Lehrkräften und die Urteile von Beobachtern herangezogen hat, konnte unter Beweis stellen, dass die Sichtweisen je nach erfasstem Merkmal teils konvergierten, teils aber auch erheblich voneinander abwichen (wenn man als Kriterium der Übereinstimmung die durchschnittliche Übereinstimmung in Form von Korrelationskoeffizienten auf Klassenebene heranzieht). Wenn Schülerangaben einer Klasse zu Mittelwerten aggregiert werden, resultieren neben zentralen Tendenzen (Mittelwert, Median etc.) auch Maße der Verteilung (wie Standardabweichung, Variationsbreite, Schiefe etc.), die sorgfältig analysiert und in theoretisch stimmiger Weise in das jeweilige theoretische Modell einbezogen werden müssen. Ob sich alle Schülerinnen und Schüler einer Klasse im Hinblick auf ein Unterrichtsmerkmal oder eine Lehrerkompetenz komplett einig sind, ob es distinkte Gruppen gibt (z.B. in Form bimodaler Verteilungen), ob Dissens oder Konsens überwiegen - dies dürfte von großer Wichtigkeit sein.

### *Motivation, Akzeptanz und lernbezogene Einstellungen*

Ebenso wenig besteht ein Zweifel daran, dass Schule und Unterricht vielfältige Ziele zum Gegenstand haben (siehe Abbildung 1), und dass dazu neben dem Aufbau intelligenten Wissens und fachlicher Kompetenzen auch die

Grundlegung lebenslangen Lernens, die Entwicklung positiver und realistischer lernbezogener Einstellungen zur Schule, zu den Lehrkräften und zum Fach gehören. Aggregiert man solche Angaben zum Unterricht, zum Lehrer-Schüler- oder Schüler-Schüler-Verhältnis auf die Klassenebene, was in der empirischen Unterrichtsforschung üblich ist, dann charakterisieren die so gebildeten Maße wichtige Eingangsvoraussetzungen für die Qualität des Unterrichts oder seine möglichen Wirkungen. Auch hier sind *Niveaumaspekte* (z.B. das Level der Lernmotivation einer Klasse oder das Angstniveau) sorgfältig von *Verteilungaspekten* zu unterscheiden (im Extremfall kann die klasseninterne Streuung gleich Null sein, die Verteilung kann bimodal, sie kann groß oder klein sein). Die Angaben einer Klasse (oder eines Kollegiums, eines Studierendenjahrgangs etc.) lassen sich je nach Urteilsgegenstand (auf den Unterricht, auf die Lehrperson, auf das Verhältnis innerhalb des Kollektivs bezogen) und nach Aggregationsebene (Individualebene, Klassen- und Schulebene) klassifizieren.

So weit, so gut. Die Klimaforschung ist ein eigener Forschungszweig und hat zu zahlreichen interessanten empirischen Ergebnissen geführt, wobei ich allerdings irritiert bin, dass ausgerechnet eines der wichtigsten deutschsprachigen Werke - das von Fend (1977) bei Eder gar nicht auftaucht. Worin liegt darüber hinaus der theoretische, praktische und methodologische Wert der Klimaforschung für die *Unterrichtsforschung*? Nach meiner Einschätzung ist das diesbezügliche Potenzial dieser Forschung zur Zeit eher gering. Ich sehe einige Schwächen dieses Ansatzes, die eine konstruktive Nutzung der hier präsentierten Erkenntnisse für die Unterrichtsforschung erschweren: Neben begrifflichen Unklarheiten und dem Verzicht auf die Nutzung moderner mehrbenenanalytischer Methoden ist dies vor allem die fehlende theoretische Grundlegung, was den Unterricht und die Lernprozesse anbelangt.

### *Theoretische und methodische Schwächen*

Ich sehe keine ausreichenden Verbindungslinien von der schulbezogenen Klimaforschung zur aktuellen psychologischen Lernforschung und vor allem nicht zur modernen Unterrichtsforschung, Instruktionspsychologie und Lernpsychologie. Eders Abbildung 1, in der die „Lernumwelt“ (Unterrichtskonzepte, Unterrichtspraktiken, Lehrerhandeln) und die „Wahrnehmung durch Betroffene“ (Schulklima, Klassenklima, Unterrichtsklima) als zwei aufeinander folgende Einheiten dargestellt werden, wird der Komplexität der Sachverhalte nicht gerecht. Das gleiche gilt für die Aussage am Ende (Abschnitt 5): „Effizienz ... und Klima .... bilden daher zwei Aspekte von Unterrichtsqualität, die grundsätzlich in einem Ursache-Wirkungs-Verhältnis zueinander stehen“. Dem würde ich widersprechen. Klima (und auch hier nur das auf den Unterricht bezogene) ist erstens kein Aspekt der Unterrichtsqualität, sondern es ist ein Reflex der Unterrichtsqualität. In meiner Abbildung 1 entspräche dies den individuellen Prozessen der Wahrnehmung und Interpretation des Unterrichts, die - auf Klassenebene gemittelt - nichts anderes als die kollektiven Wahrnehmungen sind. Zweitens stehen diese beiden Kon-



strukturiertes keineswegs in einem unidirektionalen Verhältnis zueinander. Vielmehr hat die Unterrichtsforschung gezeigt, dass die Qualität des Unterrichts und der Klassenführung selbst von vorgefundenen Rahmenbedingungen abhängt, zu denen neben dem Klassenkontext (insbesondere der Klassenzusammensetzung, z.B. Niveau und Heterogenität der Vorkenntnisse, Anteil von Schülerinnen und Schülern, für die die Unterrichtssprache eine Fremdsprache ist) selbstverständlich auch das zählt, was hier als „Klima“ bezeichnet wird (vgl. Helmke & Weinert, 1997). Das *Klassenklima* (z.B. die Wertschätzung oder auch Tabuisierung von Konformität und Leistung innerhalb der Klasse) ist eine wichtige Bedingung für den Unterricht und wird seinerseits vom Unterricht beeinflusst. Das gleiche gilt für das *Unterrichtsklima*: Der Unterricht hat massive Auswirkungen auf die Erwartungen und Sichtweisen der Schülerinnen und Schüler (auf ihre Einschätzung der Lehrperson, der didaktischen Effizienz, der Wirksamkeit der Klassenführung etc.). Diese kollektive Sicht übt (je nachdem, wie stark und wie homogen sie ist) selbst wiederum einen Einfluss auf Erwartungen, Planung und Durchführung des Unterrichts aus.

Soweit zum Unterricht. Natürlich ist die Klimaforschung auch in anderen Inhaltsbereichen verankert als dem Unterricht, so z.B. in der Entwicklungspsychologie und Sozialisationsforschung. Auch hier habe ich jedoch leise Zweifel an der theoretischen Fruchtbarkeit allgemeiner Wirkungsaussagen (so ist im Artikel von Eder die Rede davon, dass das „Klima - als ein wesentliches Merkmal von Schulen - als umfassender Prädiktor der schulischen Entwicklung von jungen Menschen gesehen“ (wird)).

In methodischer Hinsicht ist es gerade bei einem Gebiet wie der Klimaforschung, in dem es um individuelle und kollektive (geteilte) Wahrnehmungsprozesse und -ergebnisse geht, unabdingbar, mit mehrerebenenanalytischen Werkzeugen zu arbeiten. Angesichts der hierarchischen Verschachtelung (Individual-, Klassen- und Schulebene) der „Klimadaten“ sind Abschätzungen der Wirksamkeit bestimmter Variablen, die sich nur auf eine einzige Ebene beziehen (*nur* auf Individualebene oder *nur* auf Klassenebene), nicht zielführend.

### *Wie nötig ist der Klimabegriff für die Unterrichtsforschung?*

Gruehn (2000) schlägt in ihrer sehr differenzierten und detaillierten Kritik an der Klimaforschung vor, den Begriff des Klimas fallen zu lassen: „*Verzichte man auf den Begriff des Klimas zugunsten der neutraleren Bezeichnung Wahrnehmung, hätte dies den Vorteil, den Unterricht und die dort stattfindenden Interaktionsprozesse unter dem Gesichtspunkt unterschiedlicher Wahrnehmungsperspektiven der daran Beteiligten untersuchen zu können. Hiermit würde sich auch der Anschluss zu den Ergebnissen der auf Beobachtungsdaten beruhenden Unterrichts- bzw. Lehr-Lern-Forschung eröffnen*“ (S. 88f.). Diesem Vorschlag möchte ich mich anschließen. Angesichts des gegenwärtigen theoretischen und methodologischen Standes der Klimafor-

schung, so wie er sich mir in diesem Artikel darstellt, kann sich die Erforschung der Qualität des Unterrichts und seiner Wirkungen von ihr keinen großen Nutzen erwarten. Eine theoretisch befriedigende Vernetzung der Ergebnisse der unter der Rubrik „Klima“ durchgeführten, durchaus beachtlichen Untersuchungen mit dem aktuellen Stand der allgemeinen und bereichsspezifischen Unterrichtsforschung und Lernforschung wäre wünschenswert, damit die Klimaforschung hier einen Beitrag zum Erkenntnisfortschritt leisten kann. Da man Tunnel bekanntlich von zwei Seiten aus bohren kann, ist dies keineswegs nur als Kritik an der Klimaforschung, sondern ebenso als selbstkritische Aufforderung an die Unterrichtsforschung gemeint, ihrerseits theoretisch solide und begrifflich präzise Verbindungslinien zur Klimaforschung zu entwickeln.

## Literatur

- Arnold, K.-H. (in Druck). Schulentwicklung durch Rückmeldung der Lernwirksamkeit an die Einzelschule: Möglichkeiten und Grenzen der Schuleffizienzforschung. *Zeitschrift für Pädagogik*.
- Baumert, J. & Köller, O. (2000). Unterrichtsgestaltung, verständnisvolles Lernen und multiple Zielerreichung im Mathematik- und Physikunterricht der gymnasialen Oberstufe. In J. Baumert, W. Bos & R. Lehmann (Hrsg.), *TIMSS/III. Dritte Internationale Mathematik- und Naturwissenschaftsstudie. Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn* (Vol. 2, S. 271-316). Opladen: Leske + Budrich.
- Clausen, M. (2001). *Unterrichtsqualität: Eine Frage der Perspektive?* Münster: Waxmann.
- Einsiedler, W. (2001). *Ergebnisse und Probleme der Unterrichtsforschung im Primarbereich*. Nürnberg: Institut für Grundschulforschung der Universität Erlangen-Nürnberg.
- Fend, H. (1977). *Schulklima: Soziale Einflussprozesse in der Schule*. Weinheim: Beltz.
- Fend, H. (1998). *Qualität im Bildungswesen. Schulforschung zu Systembedingungen, Schulprofilen und Lehrerleistung*. Weinheim: Juventa.
- Fischler, H. (2001). Verfahren zur Erfassung von Lehrer-Vorstellungen zum Lehren und Lernen in den Naturwissenschaften. *Zeitschrift für Didaktik der Naturwissenschaften*, 7, 105-120.
- Gagné, R.M. & Driscoll, M.P. (1988). *Essentials of learning for instruction* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.
- Gruehn, S. (2000). *Unterricht und schulisches Lernen*. Münster: Waxmann.
- Helmke, A. (in Druck). *Unterrichtsqualität - Konzepte, Messung, Veränderung*. Velber: Friedrich.
- Helmke, A. & Jäger, R. S. (Hrsg.) (2002). *Die Studie MARKUS - Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontext*. Landau: Verlag Empirische Pädagogik.
- Helmke, A., Jäger, R. S., Balzer, L., Hosenfeld, I., Ridder, A. & Schrader, F.-W. (2002). *MARKUS - Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontext. Kurzbericht*. Mainz: Ministerium für Bildung, Wissenschaft und Weiterbildung.

- Helmke, A. & Schrader, F.-W. (1997). Unterrichtsbeurteilungen durch externe Beobachter. In F.E. Weinert & A. Helmke (Hrsg.), *Entwicklung im Grundschulalter* (S. 510-518). Weinheim: Psychologie Verlags Union.
- Helmke, A. & Weinert, F. E. (1997). Bedingungsfaktoren schulischer Leistungen. In F. E. Weinert (Hrsg.), *Enzyklopädie der Psychologie. Serie Pädagogische Psychologie, Psychologie des Unterrichts und der Schule* (Enzyklopädie der Psychologie, Vol. 3, S. 71-176). Göttingen: Hogrefe.
- Hosenfeld, I., Helmke, A., Ridder, A. & Schrader, F.-W. (2002). Rolle des Kontextes. In A. Helmke & R. S. Jäger (Hrsg.), *Die Studie MARKUS - Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontext. Grundlagen und Perspektiven* (S. 155-256). Landau: Verlag Empirische Pädagogik.
- Hosenfeld, I., Helmke, A., Ridder, A. & Schrader, F.-W. (2001). Eine mehrbenenanalytische Betrachtung von Schul- und Klasseneffekten. *Empirische Pädagogik*, 15 (4), 513-534.
- Köller, O., Baumert, J. & Neubrand, J. (2000). Epistemologische Überzeugungen und Fachverständnis im Mathematik- und Physikunterricht. In J. Baumert, W. Bos & R. Lehmann (Hrsg.), *TIMSS/III. Dritte Internationale Mathematik - und Naturwissenschaftsstudie. Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn* (Vol. 2, S. 229-269). Opladen: Leske + Budrich.
- Lange, H. (2001). Die bildungspolitische Bedeutung von Schulleistungsvergleichen. In G. Kaiser, N. Knoche, D. Lind & W. Zillmer (Hrsg.), *Leistungsvergleiche im Mathematikunterricht* (S. 1-28). Hildesheim: Franzbecker.
- Oser, F. (1999). „Standards in der Lehrerbildung.“ *Beiträge zur Lehrerbildung. Teil 1* 15(1), 26-37.
- Oser, F. & Oelkers, J. (Hrsg.). (2001). *Die Wirksamkeit der Lehrerbildungssysteme. Von der Allrounderbildung zur Ausbildung professioneller Standards* (Nationales Forschungsprogramm 33, Wirksamkeit unserer Bildungssysteme). Zürich: Ruegger.
- Renkl, A. & Helmke, A. (1993). Prinzip, Nutzen und Grenzen der Generalisierungstheorie. *Empirische Pädagogik*, 7, 63-85.
- Shuell, T. (1996). Teaching and learning in a classroom context. In D. Berliner & R. Calfee (Hrsg.), *Handbook of Educational Psychology* (S. 726-764). New York: Macmillan.
- Terhart, E. (2001). *Fremde Schwestern. Zum Verhältnis von Allgemeiner Didaktik und empirischer Lehr-Lern-Forschung*. Vortrag im Rahmen der 8. Fachtagung Pädagogische Psychologie in Landau, 17.-19. September 2001.
- Walberg, H. & Paik, S. (2000). *Effective educational practices*. Lausanne: International Bureau of Education.
- Wang, M. C., Haertel, G. D. & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, 63, 249-294.
- Weinert, F. E. (1998). Guter Unterricht ist ein Unterricht, in dem mehr gelernt wird als gelehrt wird. In J. Freund, H. Gruber & W. Weidinger (Hrsg.), *Guter Unterricht - Was ist das? Aspekte von Unterrichtsqualität* (S. 7-18). Wien: ÖBV Pädagogischer Verlag.
- Weinert, F. E. & Helmke, A. (Hrsg.). (1997). *Entwicklung im Grundschulalter*. Weinheim: Psychologie Verlags Union.

Anschrift des Autors:

Prof. Dr. A. Helmke

Universität Landau, Fachbereich Psychologie

Im Fort 7, 76829 Landau